MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 1, 2018/2019

DPS5018 –INTRODUCTION TO PROBABILITY AND STATISTICS

(For Diploma students only)

24 OCTOBER 2018
9.00 a.m - 11.00 a.m
(2 Hours)

INSTRUCTIONS TO STUDENT

1. This question paper consists of **8** pages (**4** pages with **5** questions and **4** pages for appendix)
2. Answer **ALL** questions. All necessary working steps must be shown.
3. Write all your answers in the answer booklet provided.
4. Statistical Tables will be provided.

QUESTION 1 [15 Marks]

(a) An established company which located at Damansara Height, Kuala Lumpur has 250 staff. 40% of them are males and the rest are females while 35 of male staff are bachelors. From all the female staff, 83 of them are married.

i. Based on information above, copy and complete the following contingency table. (4 marks)

Marital Status	Gender		
	Female	Male	Total
Bachelor			
Married			
Total			250

ii. If a staff is selected at random from the company, what is the probability that the staff is a male given that the selected staff is bachelor? (4 marks)

(b) Faculty of Information Technology purchases computers from three different companies; Hp, Lenovo and Acer to occupy the labs at 45%, 35% and 20% respectively. The percentages of defective computers from Hp, Lenovo and Acer companies are 1%, 2% and 5% respectively.

i. Construct and label the tree diagram to represent the information above. (4 marks)

ii. If a computer is selected at random, what is the probability that the computer is not defective? (3 marks)

Continued...

QUESTION 2 [25 Marks]

(a) IPS MMU recorded the study duration (in years) of a group of MMU postgraduate students who graduated in year 2017.

Duration, X	2.5	3	3.5	4	4.5	5	5.5
$P(X = x)$	$2w$	0.141	0.225	0.201	0.121	0.169	0.121

i. Show that the value of w is 0.011. (2 marks)
ii. Calculate the mean and standard deviation of the study duration. (5 marks)

(b) A study shows that 33% of the secondary students know how to send emails. If a sample of 17 students are chosen at random from a secondary school, calculate the probability of

i. exactly 3 of them know how to send emails. (3 marks)
ii. at least 10 students **do not** know how to send emails. (2 marks)

(c) It is known that most of the car owners will service their car on average of twice in a year. Determine the probability of the car owners will service their car

i. less than 3 times in a year. (2 marks)
ii. between 8 and 13 times (inclusive) in 10 years. (4 marks)

(d) The lifespan of Deka fan is normally distributed with a mean of 3 years and a standard deviation of 0.88 years.

i. What is the probability that a Deka fan chosen at random will last for more than 5.5 years? (4 marks)
ii. Determine the percentage of a selected Deka fan will only last for less than 3.75 years. (3 marks)

QUESTION 3 [10 Marks]

(a) According to a swimming coach, a group of 8 swimming athletes who has been chosen randomly will practice on average of 45 minutes and a standard deviation of 0.35 minutes per day. Construct a 98% confidence interval for the population mean of the practice duration. (4 marks)

(b) A survey reported that a sample of 285 out of 375 young adults will feel uncomfortable when going out without bring their smartphones.

i. Find the point estimate for proportion of young adults who are feeling uncomfortable. (1 mark)
ii. Construct a 92% confidence interval for the population proportion of young adult who will feel uncomfortable. (5 marks)

Continued...

QUESTION 4 [30 Marks]

(a) In June 2018, Deputy Prime Minister of Malaysia suggested to transfer 2% of husband's EPF contribution to his wife's EPF. She claimed that at least 80% of husbands will support her suggestion. Based on a sample of 240 husbands, 174 of them are strongly agreed.

i. State the null and alternative hypothesis. (2 marks)
ii. By using the hypothesis testing of one population proportion, calculate the test statistic. (2 marks)
iii. Is there any evidence to reject the claim? Test at 5% level of significance. (4 marks)

(b) According to Commissioner of Building Malacca, the mean maintenance fees charged by service apartments in Malacca is about RM220. To verify it, the project manager randomly selected a sample of 18 service apartments in Malacca and found that the sample mean of maintenance fees is RM225 with a standard deviation of RM17.3. Test at 5% level of significance. (9 marks)

(c) TM Unifi recently introduced a cheaper package with much faster speed to their users. The marketing director believes that this upgraded service would reduce the number of complaints from the users. A survey had been conducted before and after upgrading the speed to validate the claim.

	Before	After
Number of users who complained, x	66	93
Number of users being selected, n	75	125

i. Determine the sample proportion who complains TM service before and after upgrading. (2 marks)
ii. State the null and alternative hypotheses. (2 marks)
iii. By using the hypothesis testing of two population proportion, compute the test statistic, Z. (5 marks)
iv. Test the claim at 10% level of significance. (4 marks)

Continued...

QUESTION 5 [20 Marks]

(a) An insurance agent recorded the age (years) and monthly insurance premium (RM) of his potential customers as follows.

Age (years), x	Monthly Insurance Premium (RM), y
25	128
31	152
42	174
36	149
28	134
30	150
30	165
34	171
21	m
35	143

i. Given that the total monthly insurance premium for these 10 customers is RM1476. Find the value of m. (2 marks)

ii. Calculate $\sum x$ and $\sum x^2$. (2 marks)

iii. Refer to part (a) ii., show that $SS_{xx} = 317.6$. (2 marks)

iv. Given that the $SS_{yy} = 3578.4$ and $SS_{xy} = 863.8$, compute the coefficient of correlation, r. Then, comment your result. (3 marks)

v. Compute b_1 and b_0 to construct the estimated least square regression equation. (5 marks)

vi. Based on this estimation, how much that a customer needs to pay per month if she signs up an insurance plan at the age of 40 years old? (2 marks)

(b) Refer to part (a), the agent claims that there is a correlation between the age and monthly insurance premium at 10% level of significance.

i. State the null and alternative hypothesis. (2 marks)

ii. Compute the test statistic. (2 marks)

End of Page

APPENDIX – KEY FORMULA

- **Mean for ungrouped data**

$$\bar{x} = \frac{\sum x_i}{n} \quad \text{or} \quad \mu = \frac{\sum x_i}{N}$$

- **Mean for grouped data**

$$\bar{x} = \frac{\sum x_i f_i}{n} \quad \text{or} \quad \mu = \frac{\sum x_i f_i}{N}$$

Where x is the midpoint and f is the frequency of the class

- **Variance for ungrouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1} \quad \text{or} \quad s^2 = \frac{\sum x_i^{\,2} - \frac{\left(\sum x_i\right)^2}{n}}{n-1}$$

- **Variance for grouped data**

$$s^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{n-1} \quad \text{or} \quad s^2 = \frac{\sum x_i^{\,2} f_i - \frac{\left(\sum x_i f_i\right)^2}{n}}{n-1}$$

Where x is the midpoint and f is the frequency of the class

DISCRETE RANDOM VARIABLE, BINOMIAL DISTRIBUTION, AND POISSON DISTRIBUTION

- Mean of a discrete random variable x : $\mu = \sum xP(X = x)$

- Standard deviation of a discrete random variable x : $\sigma = \sqrt{\sum x^2 P(X = x) - \mu^2}$

- Binomial probability formula : $P(X = x) = \binom{n}{x} p^x q^{n-x}$

- Poisson probability formula : $P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}$

- Key formula for Binomial Probability Distribution:
  - If $p \leq 0.5$
    i. $P(X = r) = B(r) - B(r-1)$
    ii. $P(X \geq r) = 1 - B(r-1)$
    iii. $P(X \leq r) = B(r)$
    iv. $P(X > r) = 1 - B(r)$
    v. $P(X < r) = B(r-1)$
    vi. $P(a \leq X \leq b) = B(b) - B(a-1)$

  - If $p > 0.5$
    i. $P(X = r) = B(n-r) - B(n-r-1)$
    ii. $P(X \geq r) = B(n-r)$
    iii. $P(X \leq r) = 1 - B(n-r-1)$
    iv. $P(a \leq X \leq b) = B(n-a) - B(n-b-1)$

- Key formula for Poisson Probability Distribution:
  - If $\mu = \lambda$
    i. $P(X = r) = Poi(r) - Poi(r-1)$
    ii. $P(X \geq r) = 1 - Poi(r-1)$
    iii. $P(X \leq r) = Poi(r)$
    iv. $P(X > r) = 1 - Poi(r)$
    v. $P(X < r) = Poi(r-1)$
    vi. $P(a \leq X \leq b) = Poi(b) - Poi(a-1)$

NORMAL AND STANDARD NORMAL PROBABILITY DISTRIBUTION

- z-value (observed value) for an x value : $Z = \frac{x - \mu}{\sigma}$

ESTIMATION

- The $(1-\alpha)100\%$ confidence interval for μ is

$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$ where $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$

$\bar{x} \pm t_{\alpha/2,n-1}s_{\bar{x}}$ where $s_{\bar{x}} = \dfrac{s}{\sqrt{n}}$

- The $(1-\alpha)100\%$ confidence interval for p is

$\hat{p} \pm z_{\alpha/2}s_{\hat{p}}$ where $s_{\hat{p}} = \sqrt{\dfrac{\hat{p}\hat{q}}{n}}$

- **Sample size**

$$n = \left[\frac{\left(Z_{\alpha/2}\right)\sigma}{E}\right]^2 \qquad n = \frac{\left(Z_{\alpha/2}\right)^2 pq}{E^2}$$

HYPOTHESIS TESTING

- **ONE population**

Mean:

$$Z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} \quad \textbf{or} \quad t = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

Proportion:

$$Z = \frac{\hat{p}-p}{\sqrt{\dfrac{pq}{n}}}$$

- **TWO population**

Mean:

$$Z = \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^{\,2}}{n_1}+\dfrac{\sigma_2^{\,2}}{n_2}}}$$

Proportion:

$$Z = \frac{(\hat{p}_1-\hat{p}_2)-(p_1-p_2)}{\sqrt{\hat{p}_p(1-\hat{p}_p)}\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}} \qquad \text{where} \qquad \hat{p}_p = \frac{x_1+x_2}{n_1+n_2}$$

REGRESSION

- For the **least square regression equation,** $\hat{y} = b_0 + b_1 x$

  where $b_1 = \dfrac{SS_{xy}}{SS_{xx}}$ and $b_0 = \dfrac{1}{n}\left[\sum y - b_1 \sum x\right]$

- **Correlation coefficient,** r

  $$r = \frac{SS_{xy}}{\sqrt{SS_{xx} SS_{yy}}}$$

  where

  $$SS_{xx} = \sum x^2 - \frac{\left(\sum x\right)^2}{n}, \qquad SS_{yy} = \sum y^2 - \frac{\left(\sum y\right)^2}{n} \text{ and}$$

  $$SS_{xy} = \sum xy - \frac{\left(\sum x\right)\left(\sum y\right)}{n}$$

- **Regression *t*-Test**

  $$t = \frac{b_1 - \beta_1}{s_b} \qquad \text{where } s_b = \frac{s_e}{\sqrt{SS_{xx}}} \qquad \text{and } s_e = \sqrt{\frac{SS_{yy} - b_1 SS_{xy}}{n-2}}$$

- **Correlation *t*-Test**

  $$t = \frac{r}{\sqrt{\dfrac{1-r^2}{n-2}}}$$